

Politechnika Warszawska
Wydział Elektroniki i Technik Informatycznych

Warszawa, 17 października 2017 r.

D z i e k a n a t

Uprzejmie informuję, że na Wydziale Elektroniki i Technik Informatycznych Politechniki Warszawskiej odbędzie się w dniu 7 listopada 2017 r. publiczna obrona rozprawy doktorskiej

mgr inż. Łukasza Dąbały

temat: „Estimation and application of depth information in generation of view dependent effects in digital image synthesis”.

promotor – prof. dr hab. inż. Przemysław Rokita z Politechniki Warszawskiej

recenzenci:

dr hab. inż. Radosław Mantiuk, prof. Zachodniopomorskiego Uniwersytetu Technologicznego w Szczecinie

dr hab. inż. Krzysztof Murawski, prof. Wojskowej Akademii Technicznej

Obrona odbędzie się w dniu 7 listopada 2017 r. w sali 116 na Wydziale Elektroniki i Technik Informatycznych – Gmach im. Janusza Groszkowskiego, Warszawa, ul. Nowowiejska 15/19; początek godz. 9.00.

Po adresem: www.elka.pw.edu.pl/Wydzial/Rada-Wydzialu/Harmonogram-obron-doktorskich-streszczenia-i-recenzje zapewniony jest na stronie Wydziału dostęp do tekstów streszczenia rozprawy i recenzji, jak również do tekstu rozprawy umieszczonej w Bazie Wiedzy Politechniki Warszawskiej.

Dziekan



prof. dr hab. inż. Krzysztof Zaremba

Rodzaj pracy- rozprawa doktorska

Autor – mgr inż. Łukasz Dąbała

Promotor-prof. dr hab. inż. Przemysław Rokita

Tytuł pracy: "Estimation and application of depth information in generation of view dependent effects in digital image synthesis".

Streszczenie

Celem pracy była analiza dotycząca tworzenia i wykorzystania map głębokości. W tym celu stworzono algorytmy korzystające z informacji o głębi w celu zwiększenia i polepszenia informacji wizualnej prezentowanej widzowi. Stereoskopia oraz śledzenie twarzy umożliwiły stworzenie symulacji holografii na zwykłym ekranie komputera, która może być dostępna do każdego użytkownika. Głębia, jako niezależne źródło informacji została wykorzystana również podczas analizy jej wpływu na percepcję obrazów. Dzięki temu możliwe było stworzenie metody, która jest w stanie precyzyjnie przewidzieć miejsca na obrazie, które będą przyciągać uwagę widza. Dla nowej technologii light field, stworzono nowy algorytm, który jest w stanie szacować wartość głębokości dla wielu obrazów jednocześnie. Ma on tę przewagę, że informacja między poszczególnymi widokami jest uzgadniana, dzięki czemu wartości głębokości są ujednoczone we wszystkich widokach. Algorytmy zostały porównane z innymi obecnie dostępnymi rozwiązaniami.

Warszawa, 09.08.2017 r.

**RECENZJA ROZPRAWY DOKTORSKIEJ
DLA RADY WYDZIAŁU ELEKTRONIKI I TECHNIK INFORMACYJNYCH
POLITECHNIKI WARSZAWSKIEJ**

Tytuł rozprawy:

Estimation and Application of Depth Information in Generation of View Dependent effects in Digital Image Synthesis

Autor rozprawy: mgr inż. Łukasz Bernard DĄBAŁA

Druk: Drukarnia Oficyny Wydawniczej PW , Warszawa, 2017.

1. Uwagi wstępne

Recenzowana rozprawa została przedstawiona w jednym tomie składającym się z sześciu rozdziałów (w tym wprowadzenie i podsumowanie) oraz spisu publikacji autora i wykazu literatury. Tom liczy 79 stron.

2. Dziedzina badań, cel i teza rozprawy

Jakie zagadnienie naukowe jest rozpatrzone w pracy (teza rozprawy) i czy zostało one dostatecznie jasno sformułowane przez autora? Jaki charakter ma rozprawa (teoretyczny, doświadczalny, inny)?

Rozprawa dotyczy wydobywania i wykorzystywania informacji o głębi wyznaczonej na podstawie analizy obrazów cyfrowych do generowania (syntezy), efektów wizyjnych. W ramach zadań poruszanych w pracy autor rozważa wykorzystanie informacji pozyskanej w wyniku obróbki obrazu do realizacji symulowanego hologramu (rozdział 3), wykrywania sposobu postrzegania obiektów (rozdział 4) oraz realizacji techniki Light Fields. Poruszane w pracy zagadnienia od dawna są przedmiotem badań¹. Współcześnie, wraz z rozwojem architektur komputerów oraz metod

¹ 1) Thomson, LC., *Stimulation of the Retina With Light Fields of Small Size*, British Medical Bulletin, Vol. 9, 1, pp. 50-54, 1953. 2) Balakshy, *Application of acousto-optic interaction for holographic conversion of light fields*, Optics and Laser Technology, Vol. 28, 2, pp. 109-117, 1996. 3) Saykol, E., Gudukbay, U., Ulusoy, O., *Integrated querying of images by color, shape, and texture content of salient objects*, Advances in Information Systems, Proceedings, Book Series: Lecture Notes in Computer Science, Vol. 3261, pp. 363-371, 2004. 4) Yan, XY., Wang, YH., Song, Q., Dai, KH., *Salient object*

realizacji obliczeń równoległych, stały się one atrakcyjne także z uwagi na możliwość ich wykonania w stosunkowo krótkim czasie. Nie małe znaczenia ma również dostępność bibliotek², których wykorzystanie istotnie przyspiesza i upraszcza wytwarzanie oprogramowania oraz weryfikację hipotez. Mimo to, problemy występujące w obszarze przetwarzania i analizy obrazów cyfrowych są nadal bardzo aktualne. Z tego względu prowadzenie prac w tym zakresie uważam za porządane.

Cel recenzowanej rozprawy autor określił jako „*analiza dotycząca tworzenia i wykorzystania map głębi*” (str. iii). W związku z postawionym celem autor określił trzy tezy badawcze, którymi zamierzał się zająć (punkt 1.1, str. 3), w brzmieniu:

1. Usage of depth information helps to provide new experiences for the user and can be useful in terms of creation of new applications.
2. Depth as an independent source of information allows to enhance existing methods of computer vision and improve their performance.
3. Estimation of the disparity can be quick and robust for light field technology with large baselines, what makes possible to ameliorate applications for the film industry or to create new types of effects in post processing.

Formułując tezy, autor określił cel wykorzystania mapy głębi (ang. *depth map*) oraz mapy różnic (ang. *disparity map*). Wskazał, że znajomość wymienionych map może być pomocna podczas wytwarzania aplikacji komputerowych. Może również stanowić podstawę do wytwarzania bardziej wydajnych algorytmów przetwarzania i analizy obrazów.

Podsumowując tezy rozprawy są sformułowane jasno, a podjęta tematyka badawcza jest aktualna.

3. Uwagi o rozwiązaniu zadania naukowego

Czy w rozprawie przeprowadzono w sposób właściwy analizę źródeł (w tym literatury światowej, stanu wiedzy i zastosowań w przemyśle) świadczący o dostatecznej wiedzy autora. Czy wnioski z przeglądu źródeł sformułowano w sposób jasny i przekonujący?

Rozprawa wyróżnia się obszernym, 109 pozycji, wykazem literatury. W wykazie siedem prac pochodzi z Polski, w tym pięć jest autorstwa autora. W przeglądzie źródeł zwrócono uwagę zarówno

detection via boosting object-level distinctiveness and saliency refinement. Journal of Visual Communication and Image Representation, Vol. 48, pp. 224–237, 2017.

² 1) MVTec HALCON, 2) National Instruments LabVIEW, 3) MathWorks Matlab, 4) 3DExpress, 5) A&B Software's ActiveGeni, 6) Adaptive Vision Studio, 7) Aforge.NET Framework, 8) AQSENSE SAL3D, 9) Astroart, 10) Cognex VisionPro, 11) Digital Optics VPP, 12) EVT EyeVision, 13) Fastvideo, 14) ImageWarp, 15) Integrating Vision Toolkit, 16) Intopii Pinta, 17) Keyetech, 18) libdmtx, 19) Matrox MIL, 20) Micro-Manager, 21) NeuroCheck, 22) Neurotechnology SentiSight, 23) StreamPix, 24) OpenCV, 25) RoboRealm, 26) STEMMER IMAGING Common Vision Blox, 27) SANXO Modular-X, 28) Sightcorp, 29) Tordivel Scorpion, 30) VisionLab.

na sprzętowe jak i programowe wsparcie poruszanej tematyki badań. Dokonano również przeglądu literatury z zakresu ekonomii (np. pozycje [4, 6] i anatomii [10]). Siedem pozycji ze spisu literatury wydaje się być nie cytowane w pracy ([8], [58], [61], [81], [88], [92] i [103]).

Czy autor rozwiązał postawione zagadnienia, czy użył właściwej do tego metody i czy przyjęte założenia są uzasadnione?

Podstawowe zadanie badawcze polegało na analizie metod wyznaczania map głębi oraz zaproponowaniu ich wykorzystania do polepszania sposobu prezentowania informacji wizyjnej.

Zadanie pierwsze autor określił jako wytworzenie symulowanego hologramu. W tym celu połączył on techniki: widzenie stereoskopowe oraz detekcję i śledzenie twarzy, które wspierał mapą różnic (ang. *disparity map*). Uzyskane wyniki autor opublikował w dwóch publikacjach³. Dokonując analizę literatury autor rozprawy do realizacji zadania pierwszego wybrał powszechnie znany klasyfikator Viola-Jones, którego implementację można znaleźć w bibliotece Open CV. Jego zastosowanie pozwoliło stosunkowo łatwo zrealizować detekcję położenia twarzy na obrazie oraz jej elementów charakterystycznych, takich jak: oczy, nos, usta. W dalszej części pracy autor przedstawia algorytm detekcji ruchu twarzy, który jednak nie wykorzystuje cech uprzednio wyznaczonych. W zamian za to proponowany jest sposób polegający na filtracji obrazu filtrem Sobela, następnie filtrem Canny i wykonaniu operacji morfologicznych. Takie podejście jest mało zrozumiałe i wymaga uszczegółowienia, o co proszę autora w trakcie publicznej obrony. Równie niezrozumiałe jest postępowanie przedstawione w rozdziale 3 (str. 30). Autor przedstawia tam problemy wynikające z braku skutecznego wyznaczania pozycji głowy i jej błędnego śledzenia. Problemy te autor proponuje rozwiązać w naiwny sposób – wprowadzając markery umieszczone na głowie (okularach). Jest to tyle zaskakujące, że obecnie obserwowany trend rozwoju algorytmów przetwarzania i analizy obrazów główny ciężar kładzie na konstrukcję algorytmów „odpornych” (ang. *robust*).

Zadanie drugie autor określił jako opracowanie i implementację algorytmu zdolnego do wskazania obiektu „przyciągającego” uwagę (ang. *saliency*). Ta część pracy powstała na podstawie artykułu opublikowanego przez autora⁴ i stanowi powiązanie informatyki i psychologii. W pracy, jako motywację do podjęcia badań autor wskazuje brak algorytmu wyznaczania obiektu przyciągającego uwagę dającego szybko pewny wynik na bazie przetwarzania obrazu RGB (str. 36). Jak wskazuje analiza literatury tematyka ta jest aktualna do dziś. Na uzasadnienie można przytoczyć publikację⁵

³ 1) Dąbała Ł., Rokita P., *Simulated Holography Based on Stereoscopy and Face Tracking*, Springer International Publishing, pp.: 163–170, 2014. 2) Dąbała Ł., Rokita P., *Improved Simulation of Holography Based on Stereoscopy and Face Tracking*, Springer International Publishing, pp. 193–201, 2015.

⁴ Dąbała Ł., Rokita P., *Depth Guided Detection of Salient Objects*, Springer International Publishing, pp.: 197–205, 2016.

⁵ Wang S.-T., Zhou Z., Qu H.-B., Li B., *Visual Saliency Detection for RGB-D Images with Generative Model*, S.-H. Lai et al. (Eds.): ACCV 2016, Part V, LNCS 10115, pp. 20–35, 2017.

z 2016 roku, w której wykorzystano mapę głębi oraz obraz rejestrowany w przestrzeni barw RGB, tworząc model RGB-D.

W recenzowanej pracy do wyznaczenia obiektu „przyciągającego” uwagę użyto model kolorów CIELAB, który wraz z informacją o głębi tworzy model CIELABZ. Jako daną wejściową wykorzystano strukturę piramidy, reprezentującą obraz w wielu rozdzielczościach (ang. *multi-scale signal representation*). Obliczenia realizowano dla każdego poziomu piramidy. W rezultacie obliczeń określano unikalność pikseli (ang. *uniqueness*) oraz mapę odrzuceń (ang. *rejection map*). Na ich podstawie określano wynik działania algorytmu – wskazywano obiekt przyciągający uwagę.

Trzecie, ostatnie, zadanie dotyczyło wytworzenia widoku obiektu, tak by możliwa była jego prezentacja z różnych punktów widzenia (ang. *Light field*). W tym celu autor zaproponował metodę syntezy informacji pochodzącej z wielu obrazów i wielu kamer. To zadanie, zrealizowane przez autora, uważam za najtrudniejsze i najciekawsze. Mimo, iż nie podzielam opinii wyrażonej przez autora w ostatnim zdaniu na stronie 47, w którym autor twierdzi, że obecnie brak jest algorytmów wykorzystujących w tym samym czasie wielu rozdzielczości i wielu widoków obiektu (patrz przypis⁶) doceniam zakres prac wykonanych przez autora.

Na czym polega oryginalność rozprawy, co stanowi samodzielny i oryginalny dorobek autora, jaka jest pozycja rozprawy w stosunku do stanu wiedzy czy poziomu techniki reprezentowanych przez literaturę światową?

Samodzielny i oryginalny dorobek autora stanowią algorytmy do: generowania symulowanego hologramu, wyznaczenia położenia obiektu na zdjęciu, który przyciąga uwagę oraz opracowanie algorytmu generowania widoku obiektu z różnych punktów widzenia. Stanowią go również wyniki badań opracowanych algorytmów oraz sformułowane na ich podstawie wnioski. Są to także publikacje autora opublikowane w wydawnictwie Springer (3 publikacje), Computer Graphics Forum - Willey (2 publikacje). Interesującym rezultatem badań autora jest też zwrócenie uwagi na możliwość zastąpienia, w pewnych okolicznościach, mapy głębi (ang. *depth map*) przez mapę różnic (ang. *disparity map*). W tym względzie rezultaty uzyskane przez autora są zgodne z wynikami prac, które prezentowane są na konferencjach międzynarodowych. Pomysł prowadzący do zastąpienia mapy głębi przez mapę różnic umożliwił znaczne ograniczenie liczby obliczeń. Było to jednak możliwe z uwagi na założenie, że opracowane metody mają znaleźć zastosowania do generowania symulowanych widoków, a nie do realizacji rzeczywistych pomiarów.

Podjęcie badań mających na celu wykorzystanie mapy różnic i mapy głębi do wsparcia algorytmów przetwarzania i analizy obrazów cyfrowych uważam za celowe i udane. Przedstawione dla tych zagadnień algorytmy i wnioski są najważniejszym wynikiem rozprawy. Rozprawa dobrze wpisuje się w badania prowadzone obecnie na świecie.

⁶ 1) Fukushima N., Ishibashi Y., *Client Driven System of Depth Image Based Rendering*, ECTI Transactions on Computer and Information Technology (ECTI-CIT), Vol. 5, 2, pp.52 – 60, 2011. 2) Kuhn A., Hirschmüller H., Mayera H., *Multi-Resolution Range Data Fusion for Multi-View Stereo Reconstruction*, LNCS, Vol. 8142, pp. 41-50, 2013.

4. Uwagi o redakcji pracy

Czy autor wykazał umiejętność poprawnego i przekonującego przedstawienia uzyskanych przez siebie wyników (zwięzłość, jasność, poprawność redakcyjna rozprawy)?

Rozprawa napisana została w języku angielskim, w sposób nie budzący wątpliwości, że autor realizując kolejne kroki stopniowo dochodził do jej przygotowania. Rozważania zostały przedstawione w sposób poprawny, ale miejscami dość zawiły. Opracowanie zaczyna się wprowadzeniem do przedmiotu badań. Tam też krótko omówiono: układ widzenia człowieka, stereoskopię, przyczyny dyskomfortu występujące podczas oglądania symulowanego obrazu 3D, techniki wyznaczania odległości do obiektu, wyznaczanie mapy różnic (ang. *disparity map*). Następnie naszkicowano obszar, którego dotyczą rozwiązywane zadania badawcze: system wizualizacji obrazu zależny od użytkownika, wyznaczanie położenia obiektów w obrazie przykuwających uwagę, opracowanie algorytmu do syntezy obrazów umożliwiającego prezentację widoku obiektu z różnych punktów widzenia. W części pierwszej, analizie literatury, autor do pewnych kwestii podszedł zbyt ogólnikowo. Na przykład na stronie 2 pisząc: „*There are lots of algorithms, starting with cross correlation and ending with neural networks, that can estimate depth from a pair of images.*” aż się prosi wskazać przykładowe pozycje literatury, podczas gdy w pracy wątek ten pozostaje urwany. Podobne sytuacje mają miejsce również w dalszych częściach opracowania (str.: 8, 9, 10, 30, 31, 38, 46, 47, 57, 60, 69).

Moje wątpliwości i uwagi szczegółowe dotyczące pracy są następujące:

- str. 33 Podano, że „*Second, the time of the data processing must be really small, because images should be show with 60Hz for an eye, so 120Hz for a scene.*”. Moje pytanie jest następujące: jak została policzona wartość 120Hz?
- str. 5 Proszę o wyjaśnienie znaczenia pojęć „*depth map*” (str. 5) i „*disparity map*” (str. 9) oraz wskazanie tego z pojęć, które autor używa w proponowanych algorytmach?
- str. 13⁴ Proszę o sprecyzowanie pisowni „*lightfields*” czy „*Light fields*”, o którym mowa na str. 13₆?
- str. 30₁₀ Autor napisał: „*The main problem that can cause our pipeline work erroneously is a failure of the tracking or face detection method. This constitutes the heart of the system, so if something goes wrong there, then the whole system fails to work. It can be simply resolved by adding markers to track (for example glue it to the glasses). They should be easier to find and track.*”. Czy zostało to zrealizowane?
- str. 36₅ Autor napisał: „*The motivation for creation of such algorithm is that there is no such method, which can deal with RGBZ data in quick and reliable manner.*” Proszę o uzasadnienie, z jakiego powodu autor ostatecznie użył model CIELAB?

- str. 39³ Jak wyznacza się wartość funkcji $C(x,y)$ i w jaki sposób zwraca ona jednocześnie kolor i wartość odległości do obiektu (depth)?
- str. 41¹⁵ Proszę o wyjaśnienie sensu oraz podanie interpretacji wyniku dodawania wartości luminacji i głębi, o której autor napisał „*After the set of operations, both branches (luminance and depth) are combined together to create final rejection map R. To combine the maps, they are just added one to another.*”?
- str. 41 We wzorze (4.4) jak ustalana jest wartość zmiennej k .
- str. 45² Co było podstawą twierdzenia „*Comparing to others, presented methods seems to be quite successful.*”?
- str. 51₃ Proszę o ustosunkowanie się czy wzór 5.2 aby na pewno reprezentuje równanie?
- str. 53 Wzór 5.6 – czy nie powinno być $f_i^{(j)}(x, d)$ zamiast $f^{(j)}_i(x, d)$?
- str. 63₁₂ Jak policzono wskaźnik SSIM?
- str. 63₁₂ Jak rozumieć twierdzenie „*In the table 5.1 are the SSIM scores between ground truth view and re-synthesized one using the depth maps created by different algorithms.*” jeśli na str. 51 autor napisał „*What makes difference between this approach and the others is that the matching step results in creation of multi view output (the disparity map $D_i^{(j)}$) and the confidence map $C_i^{(j)}$ instead of a single depth map.*”?
- str. 70₁₆ Czy prowadząc badania dotyczące obszaru *Light fields* wykonywano kalibrację kamer oraz redukcję dystorsji obiektywów? Jeśli tak, to w jaki sposób były one prowadzane?

Uwagi natury redakcyjnej:

1. Autor konsekwentnie stosuje podpisy pod rysunkami w formacie „Figure 2.1: Tekst” podczas gdy w tekstach technicznych przyjmuje się zwykle format „Figure 2.1. Tekst”.
2. Praca zawiera nieznaczną liczbę „literówek”: str. iii⁵, str. 22¹¹, str. 36₁₄, str. 52⁵, str. 53², str. 54¹⁹

5. Uwagi krytyczne

Jakie są słabe strony rozprawy i jej główne wady?

Z lektury rozprawy można jedynie się domyślać co autor ma na myśli używając sformułowania „depth map” i „disparity map”. Miejscami można odnieść wręcz wrażenie, że stosuje on obydwa pojęcia wymiennie.

Praca stanowi złożenie trzech wątków, które wzajemnie się nie wspierają. Wprawdzie punktem wspólnym jest tu wykorzystanie informacji o głębi oraz wspólny obszar badań – analiza i

przetwarzanie obrazu, ale wydaje się, że celowym byłoby wprowadzenie większego związku pomiędzy poszczególnymi rozdziałami pracy.

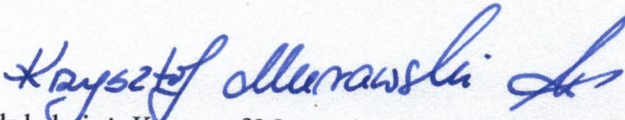
Autor skupiając się na technologii *Light fields* praktycznie pominął dwa ważne aspekty: budowę stanowiska laboratoryjnego oraz przygotowanie układu kamer do poprawnej pracy (kalibrację). Wprowadzenie tych elementów znacznie podniosło by ostateczne walory pracy.

Oceniając wyniki działania algorytmu wyznaczającego miejsca przyciągające uwagę widza (saliency) szkoda, że autor nie pokusił się o sformułowanie obiektywnych kryteriów oceny uzyskiwanych rezultatów, a opierał się jedynie na swojej własnej opinii (str. 45: „*Comparing to others, presented methods seems to be quite successful. It can handle special cases and improve the overall saliency regions detection.*”).

6. Wnioski końcowe

Na podstawie przedstawionej do oceny rozprawy doktorskiej mgr. inż. Łukasza Bernarda DĄBAŁY stwierdzam, że:

1. Recenzowana rozprawa przedstawia oryginalny sposób rozwiązania problemu naukowego. Autor wykazał, że jest uzasadnione uwzględnienie map głębi i/lub map różnic podczas wytwarzania aplikacji. Tematyka rozprawy jest aktualna.
2. Autor rozprawy wykazał się wiedzą z zakresu przetwarzania i analizy obrazów cyfrowych oraz umiejętnościami wymaganymi do rozwiązywania zagadnień naukowych i prowadzenia pracy naukowej.
3. Opiniowaną rozprawę oceniam pozytywnie: stwierdzam, że spełnia ona wymagania określone w art. 13 ust.1 ustawy z dnia 14 marca 2003 r. o stopniach naukowych i tytule naukowym oraz o stopniach i tytule w zakresie sztuki (Dz.U. z 2003 r. Nr 65, poz. 595 z późn. zm). Wnioskuje o dopuszczenie rozprawy doktorskiej do publicznej obrony.


dr hab. inż. Krzysztof Murawski, prof. WAT

dr hab. inż. Radosław Mantiuk, prof. ZUT
Wydział Informatyki
Zachodniopomorski Uniwersytet Technologiczny w Szczecinie

KWESTIONARIUSZ - RECENZJA ROZPRAWY DOKTORSKIEJ DLA RADY WYDZIAŁU
ELEKTRONIKI I TECHNIK INFORMACYJNYCH POLITECHNIKI WARSZAWSKIEJ

Tytuł rozprawy:

Estimation and Application of Depth Information in Generation of View Dependent Effects in Digital Image Synthesis

Autor rozprawy:

Łukasz Dąbała

1. Jakie zagadnienie naukowe jest rozpatrzone w pracy (teza rozprawy) i czy zostało ono dostatecznie jasno sformułowane przez autora? Jaki charakter ma rozprawa (teoretyczny, doświadczalny, inny)?

Tematem rozprawy są zagadnienia związane z estymacją i wykorzystaniem informacji o głębi w obrazowaniu komputerowym. Wspomniana informacja o głębi określa odległość od obserwatora do obiektów prezentowanych na obrazie, dodając w ten sposób trzeci wymiar do dwuwymiarowych obrazów komputerowych. W rozprawie proponowana jest metoda estymacji głębi w postaci mapy głębokości na podstawie pola światła (ang. light field), tzn. szeregu obrazów sceny wykonanych dla różnych ujęć kamery. Wyznaczona mapa głębokości wykorzystana jest do obliczenia różnicy głębokości dla obrazowania stereoskopowego, symulacji efektu głębi ostrości czy generowania obrazów za pomocą techniki chmury punktów (ang. point cloud). Kolejnym zagadnieniem poruszonym w rozprawie jest wyświetlanie obrazów holograficznych poprzez analizę kierunku patrzenia obserwatora estymowanego na podstawie położenia jego głowy. W rozprawie poruszany jest również problem estymacji obszaru uwagi wzrokowej obserwatora patrzącego na scenę. Do polepszenia jakości estymacji wykorzystana została informacja o głębi.

Autor rozprawy sformułował trzy tezy, które w przybliżonym tłumaczeniu z języka angielskiego brzmią:

- Użycie informacji o głębokości może być użyteczne w wielu aplikacjach.
- Głębokość jest niezależnym źródłem informacji, która pozwala ulepszyć istniejące metody widzenia komputerowego oraz zwiększyć ich wydajność.
- Na podstawie pól światła można wydajnie i powtarzalnie estymować różnicę głębokości (ang. disparity) co umożliwia zastosowanie generowanych wyników w przemyśle filmowym oraz do uzyskiwania nowych efektów na etapie końcowej obróbki obrazów.

Tezy sformułowane są w sposób jasny i dotyczą aktualnych i ważnych z naukowego punktu widzenia zagadnień związanych z wykorzystaniem informacji o głębi w obrazowaniu komputerowym.

Autor postawił sobie za cel rozwiązanie problemów:

- Zidentyfikowania obszarów obrazowania komputerowego, w których istotna jest informacja o głębi.
- Opracowania aplikacji symulujących wyświetlacz holograficzny poprzez śledzenie położenia

- głowy i stereoskopowe wyświetlanie obrazu.
- Opracowania systemu estymującego informację o głębokości na podstawie pola światła.
- Opracowanie algorytmu do estymacji obszarów uwagi z wykorzystaniem informacji o głębokości.

Wymienione problemy mają naukowy charakter. Ich rozwiązanie nie jest trywialne i przyczyni się do rozwoju dyscypliny naukowej obrazowania komputerowego, w szczególności widzenia komputerowego.

Rozprawa ma charakter eksperymentalno-analityczny. Opracowane oprogramowanie oraz jego testy doprowadziły do sformułowania ogólnych wniosków o możliwości estymacji głębi na podstawie pól światła, wykorzystania śledzenia twarzy do wyświetlania holografii czy skuteczności wykorzystania informacji o głębi do estymacji uwagi wzrokowej człowieka.

2. Czy w rozprawie przeprowadzono w sposób właściwy analizę źródeł / w tym literatury światowej, stanu wiedzy i zastosowań w przemyśle / świadczący o dostatecznej wiedzy autora. Czy wnioski z przeglądu źródeł sformułowano w sposób jasny i przekonujący?

Tematyka rozprawy dotyczy aktualnych problemów, które wymagają rozwiązania i znajdują się w kręgu zainteresowań środowiska akademickiego. Autor zamieszcza w rozprawie liczne odwołania do literatury naukowej (109 pozycji w tym 5 własnych) w odpowiedni sposób analizując obecny stan wiedzy związany z tematyką rozprawy.

Autor w skrótovej formie wprowadza w zagadnienia sposobu działania układu wzrokowego człowieka, widzenia stereoskopowego, sposobów estymacji głębi oraz różnicy głębokości, uwagi wzrokowej oraz pól światła. Liczne są nawiązania do literatury naukowej w rozdziałach opisujących zrealizowane cele pracy. Świadczy to o szczegółowej znajomości dokonań naukowych w konkretnych rozwiązaniach.

3. Czy autor rozwiązał postawione zagadnienia, czy użył właściwej do tego metody i czy przyjęte założenia są uzasadnione?

Na największą uwagę w rozprawie zasługuje rozdział, w którym zaproponowano algorytm estymacji głębi na podstawie pól światła. Działanie algorytmu oparte zostało na uogólnieniu techniki przepływu optycznego (ang. optical flow) Lucas-Kanade na pole światła reprezentowane przez macierz obrazów. W przyjętym rozwiązaniu uwzględniana jest wzajemna zależność pomiędzy obrazami w macierzy. Technika rozszerzona została o etap konsolidacji, w którym odrzucane są zaszumione informacje. W zaproponowanym algorytmie możliwe jest wyznaczenie różnicy głębokości pomiędzy poszczególnymi widokami w polu światła oraz mapy głębokości dla poszczególnych widoków. Algorytm zaimplementowany został na procesorze graficznym (GPU) z wykorzystaniem piramidy obrazów o różnej rozdzielczości (ang. multi-resolution representation) co umożliwi wykorzystanie go do interaktywnej obróbki pól światła.

Poprawność działania algorytmu udowodniona została na przykładzie aplikacji symulującej głębię ostrości poprzez wykorzystanie mapy głębokości. Zaproponowano system do renderingu chmury punktów, w którym do generowania obrazu wykorzystywane są piksele obrazów pola światła oraz mapa głębokości. Proponowany algorytm porównany został do technik SGM, CVF oraz Wang i inni. W szczególności ocenie poddano mapy różnic głębokości generowane przez poszczególne techniki.

Uzyskane wyniki wskazują na to, że do estymacji głębi na podstawie pól światła użyte zostały prawidłowe metody, które doprowadziły do konkretnych rezultatów aplikacyjnych.

Autor zaproponował system do wyświetlania obrazów holograficznych, w którym treść wyświetlanego obrazu zmienia się wraz z ruchem głowy obserwatora. Położenie głowy estymowane jest poprzez wykrywanie i śledzenie twarzy za pomocą obliczania przepływu optycznego na zdjęciach

tworzy rejestrowanych przez kamerę. Technika uwzględnia ruch głowy w trzech kierunkach względem wyświetlanego obrazu oraz skręcenie głowy. Obraz holograficzny wyświetlany jest w postaci obrazu anaglifowego lub za pomocą okularów polaryzacyjnych.

Przyjęte rozwiązania są poprawne i wskazują na poprawne działanie systemu. Analiza rezultatów sprowadzona została do omówienia wad i zalet działania systemu oraz do jego subiektywnej oceny.

W kolejnym rozwiązaniu Autor wykorzystuje informację o głębi do poprawy estymacji uwagi wzrokowej człowieka. Opisywana metoda bazuje na pracy [Perazzi i inni, 2012] rozszerzając prezentowaną w niej technikę o uwzględnienie głębi oraz usuwając z analizy zaszumione obszary. Pierwotna metoda bazuje na analizie zmian kontrastu występujących w obrazie. Proponowane rozszerzenie uwzględnia również zmianę głębokości w sąsiedztwie analizowanego piksela. Technika opracowana przez Autora rozprawy zestawiona została z mapami uwagi obliczonymi przez szereg innych algorytmów (głównie takich, które nie wykorzystują informacji o głębokości), wskazując na zbieżność rezultatów i poprawność przyjętych założeń.

4. Na czym polega oryginalność rozprawy, co stanowi samodzielny i oryginalny dorobek autora, jaka jest pozycja rozprawy w stosunku do stanu wiedzy czy poziomu techniki reprezentowanych przez literaturę światową?

Autor zaproponował technikę obliczania głębi dla widoków pola światła. Wspomniana technika, przy zachowaniu dokładności estymacji głębi jest dużo szybsza od rozwiązań opisywanych w literaturze umożliwiając interaktywną edycję pól światła. Każda z przedstawionych aplikacji techniki jest oryginalnym rozwiązaniem, które wskazuje na poprawność estymacji głębi na potrzeby symulacji głębi ostrości, renderingu chmury punktów i estymacji różnicy głębokości na potrzeby wyświetlaczy stereoskopowych.

Oryginalnym rozwiązaniem jest rozszerzenie estymatora uwagi wzrokowej człowieka. Wykorzystanie głębi jako dodatkowej wskazówki uwagi wspomagającej analizę kontrastów przyczyniło się do poprawy dokładności estymatora.

Rezultaty opisywane w rozprawie świadczą o wysokim poziomie naukowym i technicznym opracowanych rozwiązań. Autor wykazuje się umiejętnościami naukowego rozwiązywania problemów oraz znaczącymi kompetencjami technicznymi.

5. Czy autor wykazał umiejętność poprawnego i przekonującego przedstawienia uzyskanych przez siebie wyników / zwięzłość, jasność, poprawność redakcyjna rozprawy/?

Praca napisana została w języku angielskim, który nie jest ojczystym językiem autora. Zawiera stosunkowo liczne potknięcia językowe, które nie utrudniają jednak zrozumienia treści rozprawy.

Układ pracy jest czytelny i zwięzły, narzucony przez publikacje, z którymi skojarzone są kolejne rozdziały. Treść rozdziałów trzeciego, czwartego oraz w szczególności piątego mogłaby zostać rozszerzona. Akceptuję jednak formę pracy będącą streszczeniem treści artykułów, do których nawiązuje. Uważam, że praca zyskałaby, gdyby wspomniane rozdziały zastąpione zostały dosłowną treścią wspomnianych artykułów.

W pracy natknąłem się na bardzo nieliczne błędy redakcyjne (np. "... operation. fig 5.6 ..." -> "... operation. Fig 5.6 ...", rysunki 5.10 i 5.11 powinny zostać obrócone o 90 stopni). Mam wątpliwości co do potrzeby zamieszczenia prostych schematów, np. rys. 3.2, 3.3, 4.6.

6. Jakie są słabe strony rozprawy i jej główne wady?

Praca po uwzględnieniu skojarzonych z nią artykułów nie zawiera istotnych wad ani

znaczących uchybień. Nasuwające się spostrzeżenia mają charakter dyskusyjny i nie podważają mojej pozytywnej opinii o całości rozprawy.

Autor rozprawy uznaje możliwość generowania mapy głębokości dla każdego widoku w polu światła za podstawą cechę i zaletę opisywanego w rozprawie algorytmu. W treści rozprawy nie ma jednak wyraźnego wskazania aplikacji, w których takie rozwiązanie jest kluczowe i przewyższa pod względem jakościowym rozwiązanie z jedną mapą głębokości.

Poprawność działania systemu do wyświetlania holografii oceniona jest w sposób subiektywny. Interesujące byłoby wskazanie przez Autora w jaki sposób można weryfikować poprawność działania takiego systemu. Jakie testy/eksperymenty należałoby wykonać, aby ocenić wpływ niedokładności estymacji położenia głowy na poprawność wyświetlania holografii?

Rezultaty działania estymatora obszarów uwagi wzrokowej porównane są na zasadzie zestawienia wyników z wynikami obliczonymi przez inne techniki. Interesujące byłoby porównanie wyników za pomocą metryki obiektywnej (np. SSIM). W rozprawie nie jest opisane dla jakich parametrów wygenerowane zostały wyniki z rys. 4.11. Moim zdaniem do oceny proponowanego algorytmu niezbędne są dane referencyjne, które dodatkowo pozwoliłyby na kalibrację algorytmu (dobór najlepszych parametrów).

7. Jaka jest przydatność rozprawy do nauk technicznych?

Poprawna estymacja obszaru uwagi wzrokowej jest istotnym i nierozwiązanym do tej pory problemem naukowym. Włączenie do analizy informacji o głębi powinno przyczynić się do poprawy jakości estymatorów, ponieważ układ wzrokowy człowieka wykorzystuje dane o trzecim wymiarze do analizy obrazów.

Moim zdaniem rejestracja i przetwarzanie pól światła jest obecnie najważniejszym problemem naukowym fotografii obliczeniowej oraz technik widzenia komputerowego. Z uwagi na ograniczenia innych rozwiązań technicznych rejestrujących głębię (np. kamer typu time-of-flight), rejestracja pól światła może stać się efektywniejszym i docelowym rozwiązaniem. Stąd wysoko oceniam prezentowaną w pracy technikę szybkiego obliczania map głębokości na podstawie widoków pola światła.

8. Do której z następujących kategorii Recenzent zalicza rozprawę:

- a/ nie spełniająca wymagań stawianych rozprawom doktorskim przez obowiązujące przepisy**
- b/ wymagająca wprowadzenia poprawek i ponownego recenzowania**
- c/ spełniająca wymagania**
- d/ spełniająca wymagania z wyraźnym nadmiarem**
- e/ wybitnie dobra, zasługująca na wyróżnienie**

Stwierdzam, że Autor wniósł istotny wkład do dyscypliny naukowej informatyka w zakresie grafiki komputerowej oraz widzenia komputerowego.

Uważam, że przedłożona do recenzji rozprawa spełnia z nadmiarem wymagania Ustawy o stopniach naukowych i tytule naukowym z 14 marca 2003 roku, Dziennik Ustaw Nr 65, poz. 595 (z późniejszymi zmianami). Biorąc pod uwagę fakt, że Doktorant jest autorem dwóch artykułów publikowanych w czasopiśmie z listy A MNiSW, prezentowanych na uznanych konferencjach naukowych z dziedziny grafiki komputerowej (Eurographics'14 i Pacific Graphics'16), uważam, że praca mogłaby zostać wyróżniona.

